CRD

COMPUTATIONAL RESEARCH DIVISION

Grid and P2P Research Experiences

Karlo Berket

Distributed Systems Department, LBNL

March 12, 2004

Office of Science
U.S. DEPARTMENT OF ENERGY

Outline

- DSD Research Overview (brief)
- Reliable and Secure Group Communication
- Scalable and Secure P2P Information Sharing
  - Security challenges
  - Scalability challenges
- Grid and P2P Convergence?

DSD Research Overview

- Secure Grid Technologies
  - DOE Science Grid
  - CogKit
    - pyGlobus
  - Grid Services
    - pyGridWare
  - Distributed authorization
    - Akenti

DSD Research Overview: Data Intensive Distributed Computing

Application Events

Copy output data

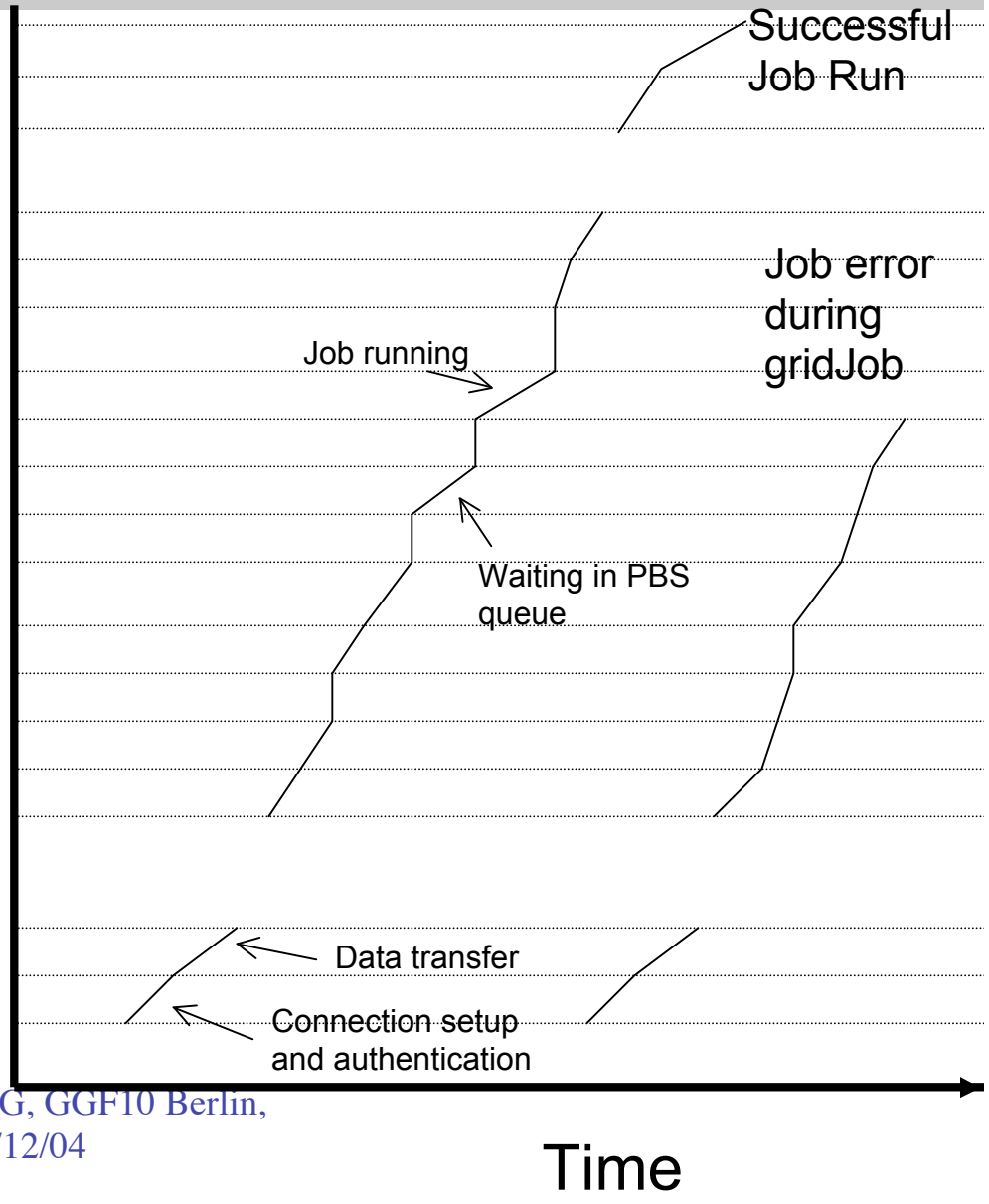
GlobusUrlCopy.put.end
GlobusUrlCopy.put.transferStart
GlobusUrlCopy.put.start

Run Grid Job

GlobusJobRun.end
jobManager.end
jobManger.jobState.done
gridJob.end
gridJob.start
jobManager.jobState.active
jobManager.jobState.pending
akentiAuthorization.end
akentiAuthorization.start
gateKeeper.end
jobManager.start
gateKeeper.start
GlobusJobRun.start

Copy input data

GlobusUrlCopy.get.end
GlobusUrlCopy.get.transferStart
GlobusUrlCopy.get.start



DSD Research Overview

- Collaboration Technologies
  - Pervasive Collaborative Computing Environment
    - IM
    - workflow
  - Reliable and Secure Group Communication
    - InterGroup
    - Secure Group Layer (SGL)
  - Scalable and Secure P2P Information Sharing
    - scishare
    - firefish

Outline

- DSD Research Overview (brief)
- Reliable and Secure Group Communication
- Scalable and Secure P2P Information Sharing
  - Security challenges
  - Scalability challenges
- Grid and P2P Convergence?

CRD Reliable and Secure Group Communication

- Provide efficient, reliable, and secure communication between collaborating sites
- Multicast communication channel directly connecting the participants
- Support participants spread across the Internet
- Support ad hoc formation of groups
- Remove dependence on servers

InterGroup Goals

- Support a broad range of applications
  - Provide a variety of guarantees
    - Reliable and unreliable delivery
    - Sender order, total order, and unordered
- Scale to the Internet
  - Split group into a sender and receiver group

InterGroup Protocols Scaling

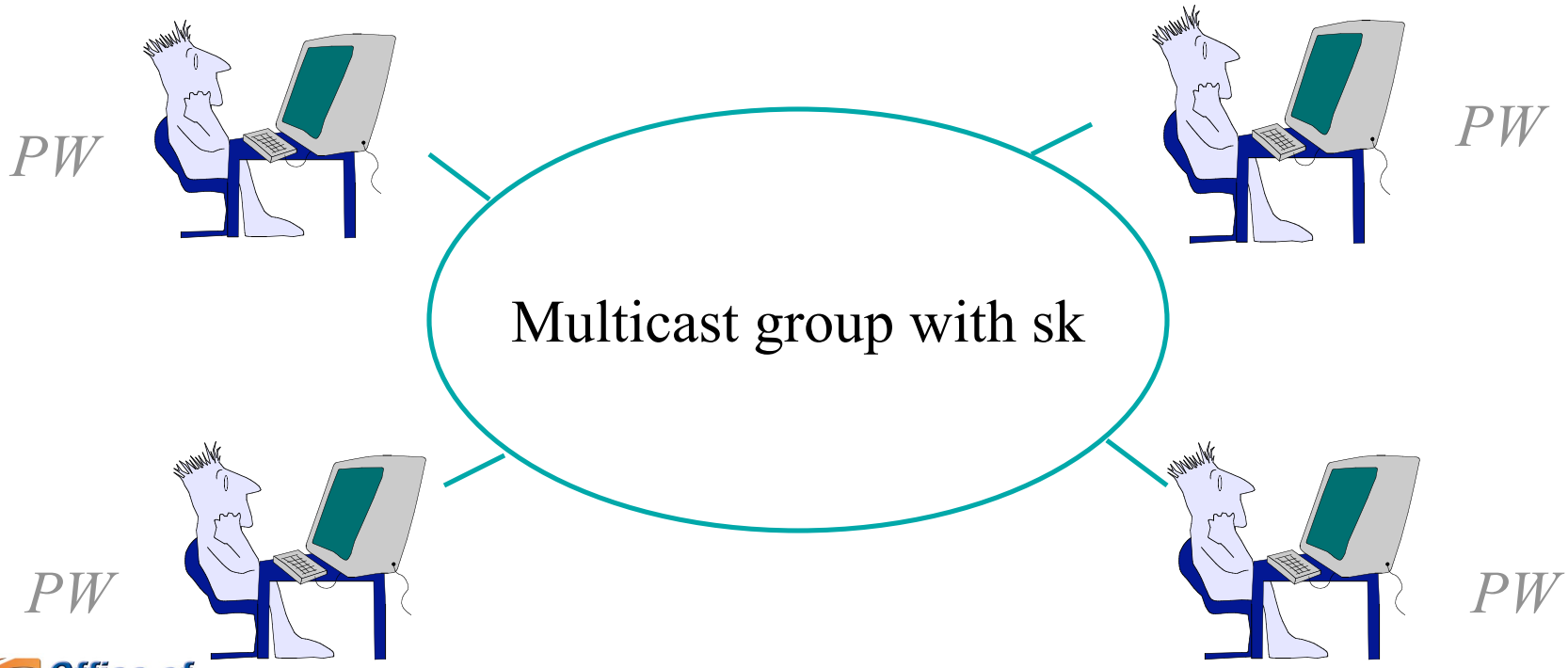
- Sender group membership
  - processes are in the sender group only while transmitting messages
  - strictly maintained
  - very dynamic
- Receiver group membership
  - not strictly maintained
  - hierarchically organized to scale to large groups
  - used for retransmissions and garbage collection

Secure Group Layer (SGL) Goals

- Provide a secure channel for the group with properties similar to SSL
- Group authorization and access control is individually enforced
- Fully distributed group key management (not centralized)
- Portable implementation

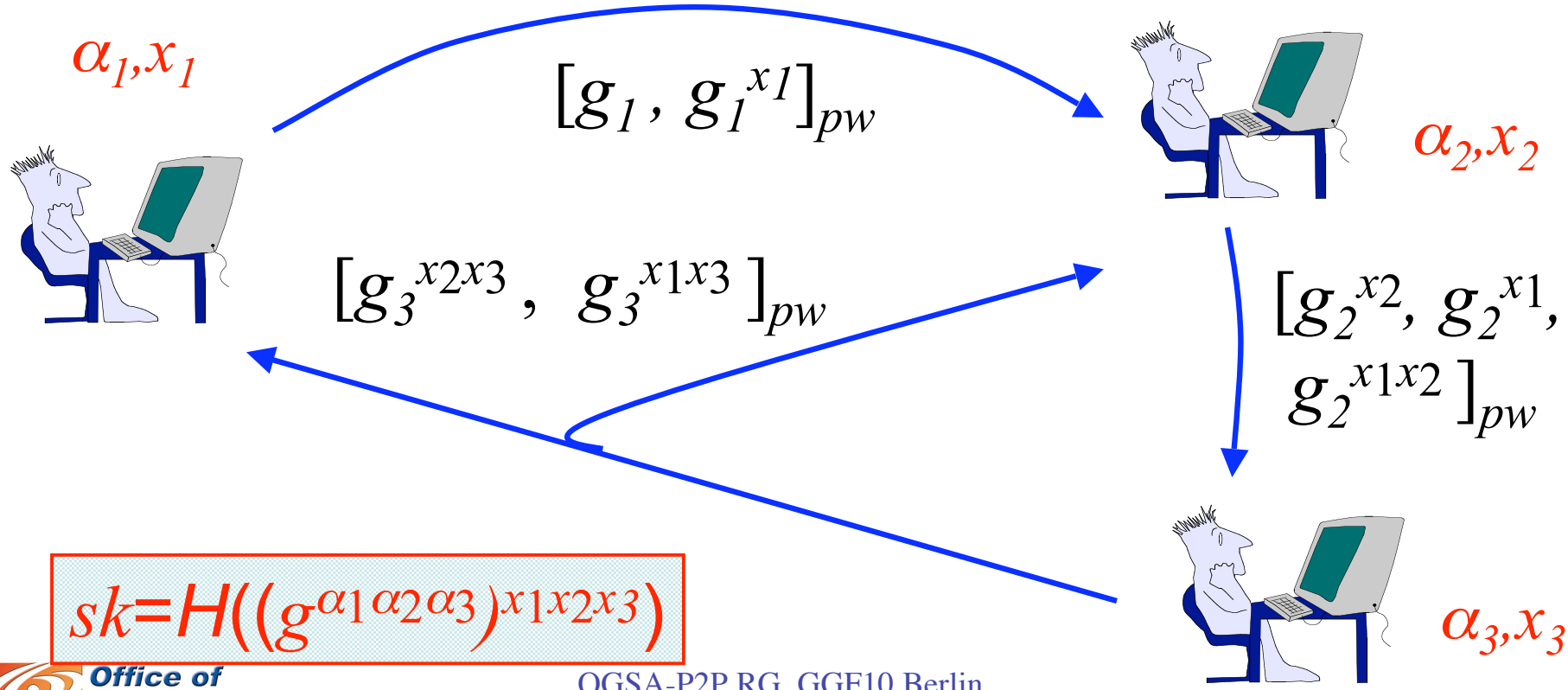
SGL Model of Communication

- A set of *n* players
  - each player is represented by an oracle
  - each player holds a low-entropy secret (PW)
- A multicast group consisting of a set of players



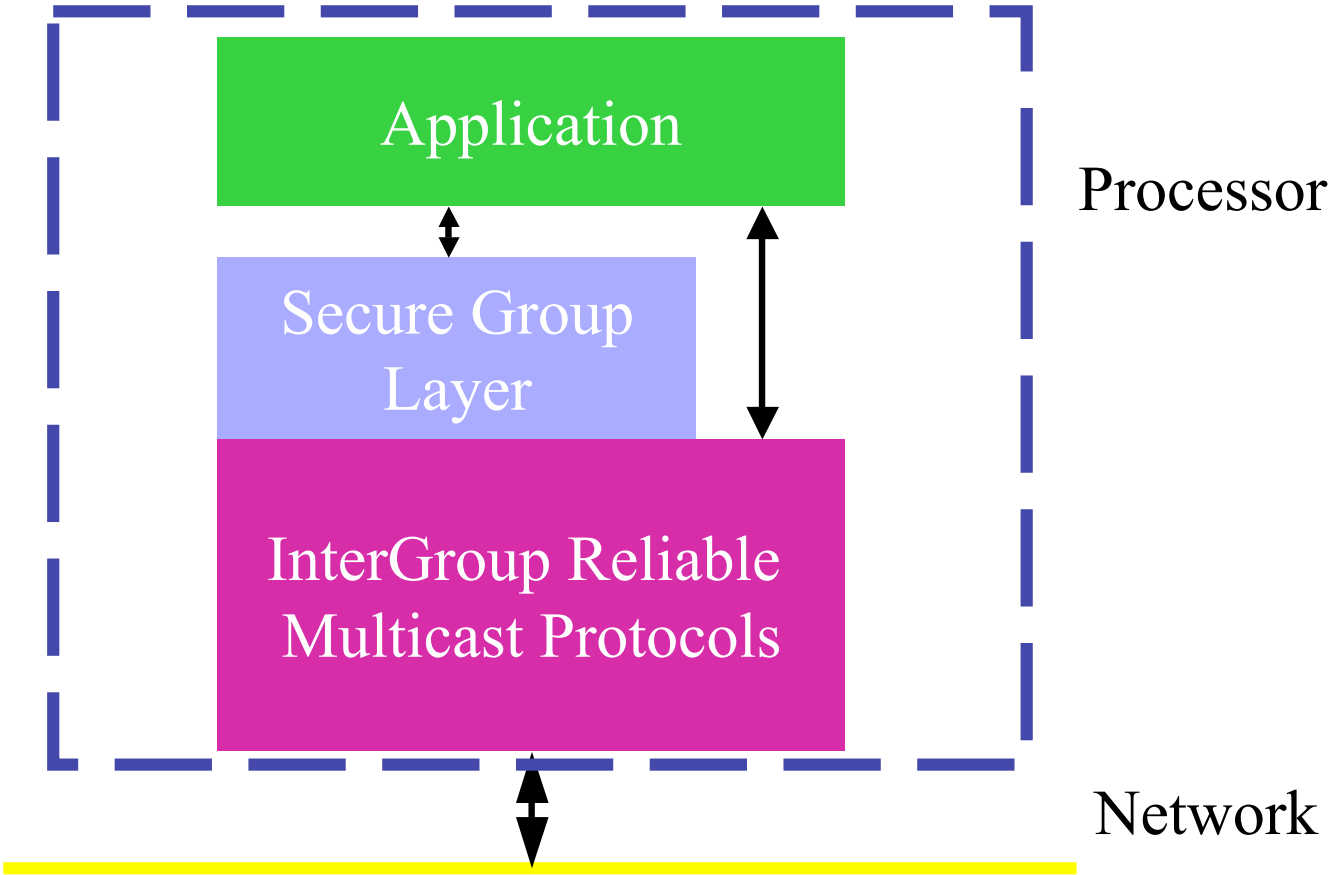
Group Diffie-Hellman Algorithm

- **Up-flow: U_i raises received values to the power of the values (x_i,α_i) and forwards to U_{i+1}**
- **Down-flow: U_n processes the last up-flow and broadcasts**



$$sk=H((g^{\alpha_1\alpha_2\alpha_3})^{x_1x_2x_3})$$

InterGroup + SGL



Status

- InterGroup
  - Node software v1.5 (Java)
  - Client software v1.5 (C++, Java, Python)
- SGL
  - Prototype using InterGroup in testing (C++)

Outline

- DSD Research Overview (brief)
- Reliable and Secure Group Communication
- Scalable and Secure P2P Information Sharing
  - Security challenges
  - Scalability challenges
- Grid and P2P Convergence?

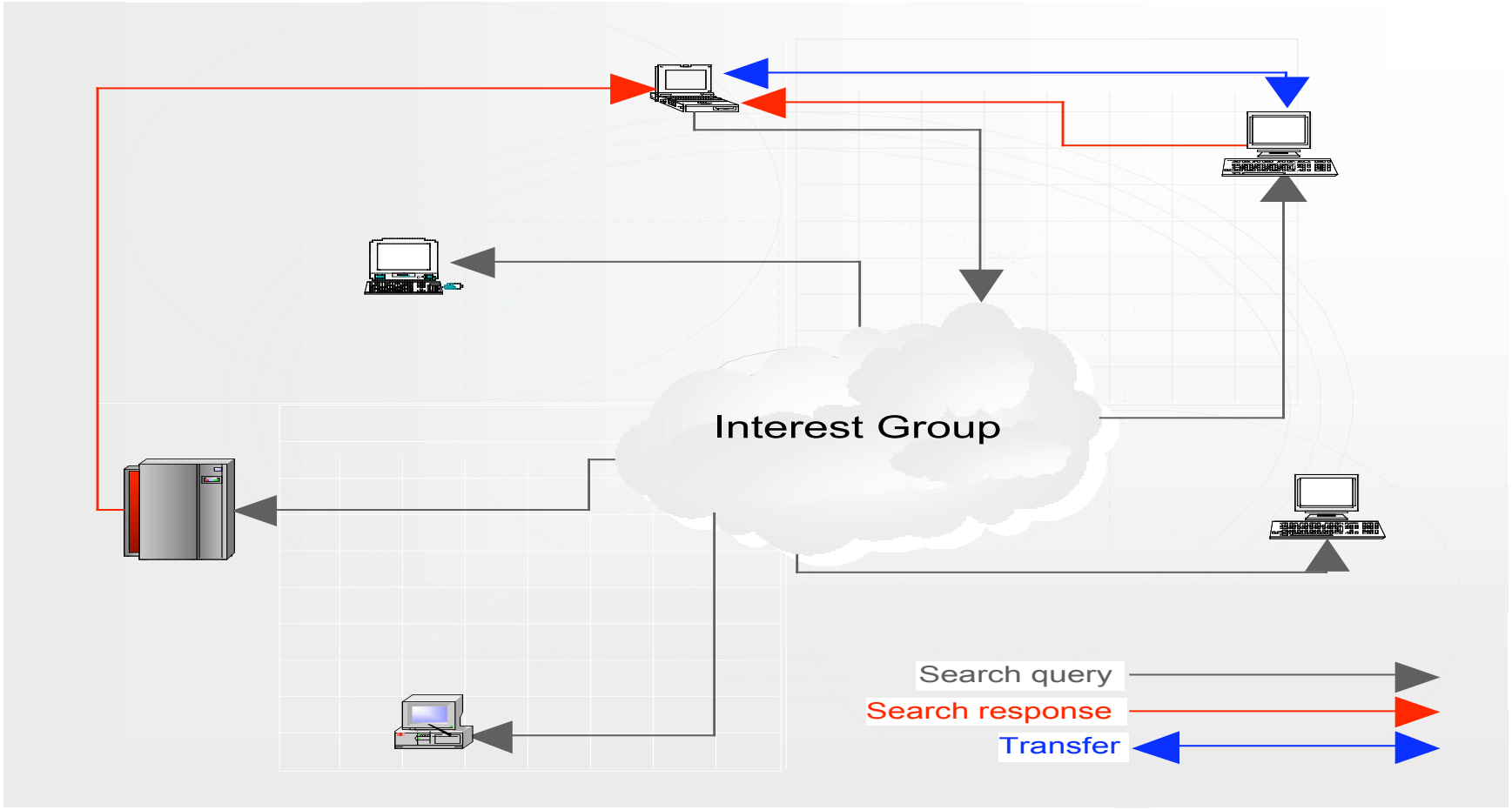
Scalable and Secure P2P Information Sharing

- Create a peer-to-peer system to support location independent information sharing in the scientific community
- Goals
  - Security
  - Scalability

Security Goals

- Confidentiality and integrity of communication
- Fine-grained access control to resources (files)
- Assumption
  - X.509 identity certificates

Communication Patterns



Securing Communication

- Search query
  - Multicast
  - SGL (work in progress)
- Search response and transfer
  - Unicast
  - HTTPS (implemented)

Access Control Goal

- Owner of resource has complete control over controlling access to that resource
  - May grant rights to control access to third parties

Distributed Authorization - Akenti

- Target widely distributed environments
  - Resources (instruments, executables, …)
  - Principals:
    - Resource owners (stakeholders)
    - User-Attribute Issuers
    - Users
- Collaborative/Grid environments that could span many autonomous/dispersed organizations.
- Provide a flexible and secure way for stakeholders to remotely and independently define authorization policy and allow fine-grained access control.

Akenti Uses 5 Types of Signed Certificates

- Policy Certificates – define trust relationships
  - Who is trusted to issue UseCondition certificates
  - What CA's are trusted to issue X.509 identity certificates
  - Where certificates can be found
- Use-Condition Certificates - express access control info
  - Contain rules for granting access to a resource
  - Can apply to one or a set of resources
  - Rights from multiple UseConditions for the same resource are additive
- Attribute Certificates – define a characteristic of a user
  - Access to resources can be based on the attributes of a user
  - Issued by trusted attribute authorities
- X.509 public key certificates – define an identity
  - Standard PKI certificate
  - Issued by a trusted Certificate Authority (CA)
- Capability Certificates –express an authorization decision
  - Grant specific rights to a resource for a user

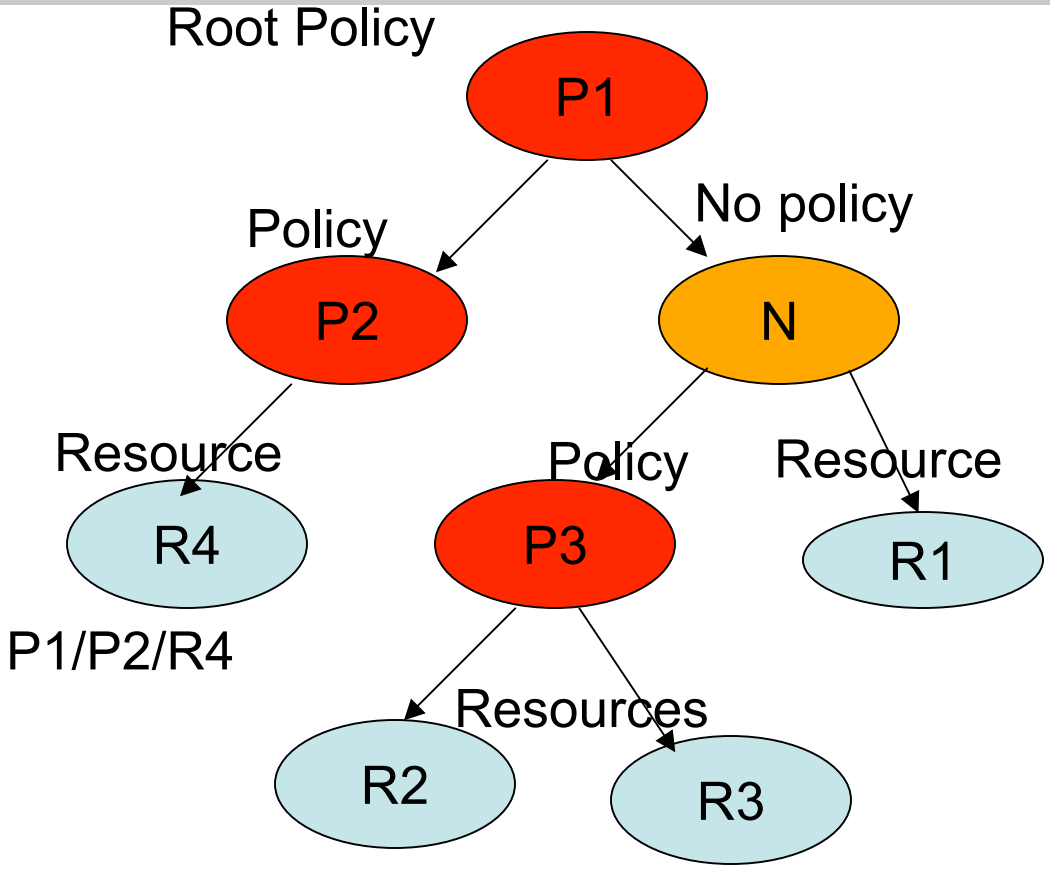
Hierarchical Resource/Policy Model

Promotes policy inheritance

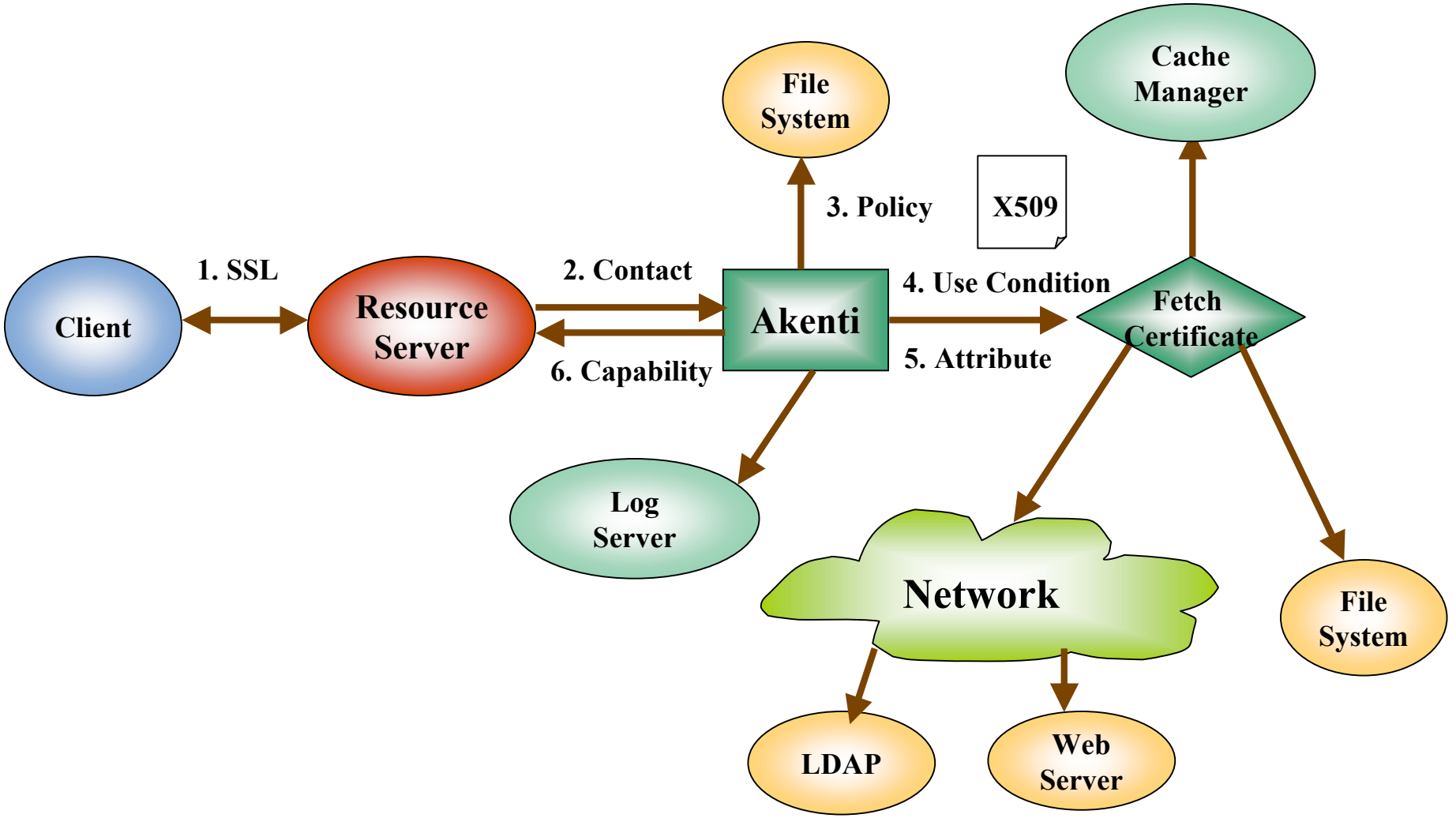
Global policies/constraints should be specified at the top level

Resources are easily aggregated

Mappings allow flexible resource names (/bin/ls, miscroscope)



Pull Model Architecture



Modifying Akenti for P2P

- Avoid incorrect denial of service
  - Akenti instance at every peer
- Simplify policy creation
  - Automate much of policy creation
  - User and group based access control
  - Make policies easily reusable

User and Group Based Access Control

- Create groups of users
- Simple Use Conditions
  - Allow all authenticated (root policy CAs)
  - Allow these groups
  - Deny these groups
  - Deny these users
  - Deny all
- Example:

  Allow all authenticated users except for group OGSA and user Karlo

Policy Reuse

- Policies created independently of resources
  - Apply policy to multiple independent resources
  - Policies can be shared among collaborators

Access Control Status

- Integrated component of scishare
  - Prototype demo at SC 2003
  - Release planned for April 2004
- Working on
  - Scishare independent version

Initial scishare Scalability Goal

- Reduce average bandwidth across all peers
- Analysis
  - Best decentralized method:
    - Multicast transport request
    - Unicast response

Resource Discovery Messaging Framework (RDMF)

- Application level messaging protocol
  - Allows for opaque queries and results
  - One-hop requests
- Implementation
  - InterGroup/SGL for distributing queries
  - HTTP(S) for responses

Discovery Use Cases

- **Within RDMF scope**
  - **Find all (available) services with > 10 MB/s**
- **Outside RMDF scope (inefficient at best)**
  - **Return the number of replica catalog services**
  - **Find the service with the largest uptime**
  - **Find all (execution service, storage service) pairs where both services of pair live in the same domain**

P2P Resource Discovery Infrastructure Requirements

- **Generality and Extensibility**
  - Application-specific multi-hop queries
  - Simple to complex queries
  - Smart dynamic routing
  - Transport independent
- **Scalability and Reliability**
  - Reduce network and client burden
  - Familiar & scalable I/O abstractions
  - Routed and direct response modes

Background

- RDMF
- Unified Peer-to-Peer Database Framework (UPDF)
- Peer Database Protocol (PDP)

W. Hoschek, "Peer-to-Peer Grid Databases for Web Service Discovery", *Grid Computing: Making the Global Infrastructure a Reality*, Editors: Fran Berman, Geoffrey Fox and Tony Hey, November 2002, Wiley Press.

Peer-to-Peer I/O (P2PIO)

- Application level messaging framework
- Supports requirements for P2P resource discovery infrastructure

K. Berket, A. Essiari, D. Gunter, W. Hoschek, "Peer-to-Peer I/O (P2PIO) Protocol Specification", http://dsd.lbl.gov/firefish/p2pio-spec/spec.pdf (work in progress)

P2PIO Mechanisms

- Opaque queries and results
- Supports iterative retrieval of result set
- Supports client choice of response mode
  - Routed
  - Direct with invitation
- Supports push and pull result retrieval
- Support client ability to influence routing
  - Neighbor selection query

P2PIO Queries

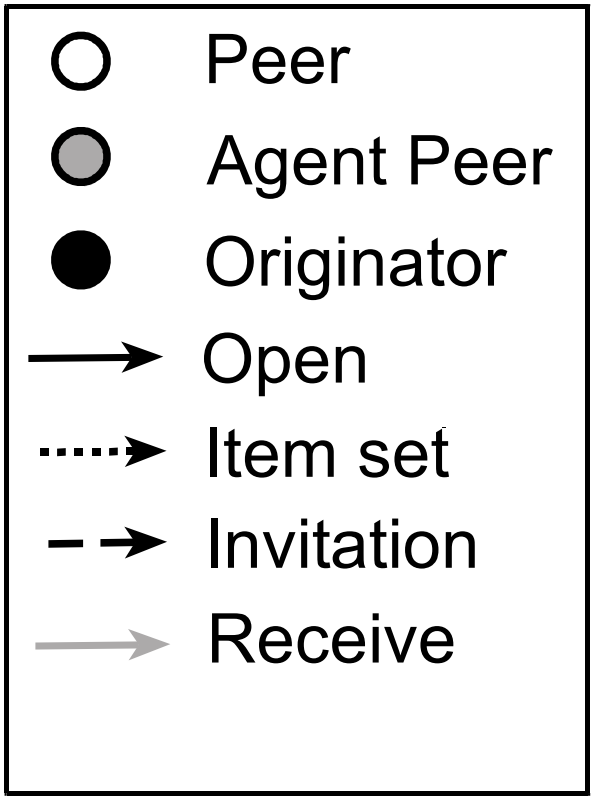
- Opaque queries and results
- Simple query (recursively partitionable)
  - Find all (available) services with > 10 MB/s
- Medium query (recursively partitionable)
  - Return the number of replica catalog services
  - Find the service with the largest uptime
- Complex query (not recursively partitionable)
  - Find all (execution service, storage service) pairs where both services of pair live in the same domain

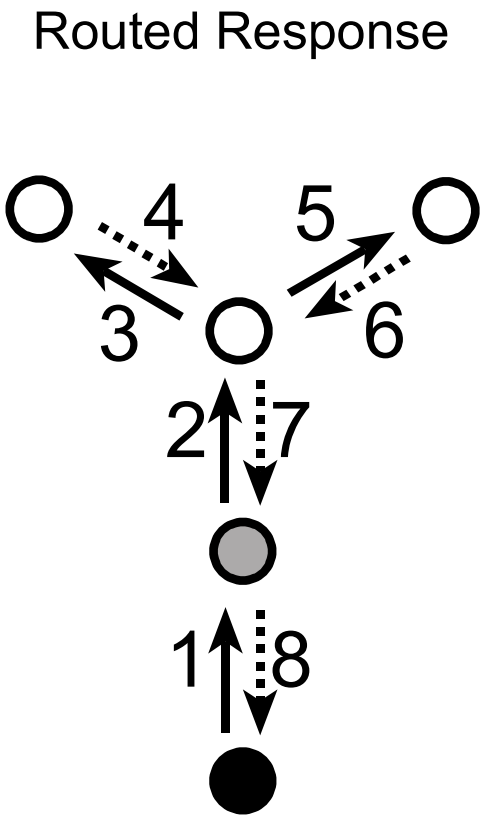
P2PIO Abstractions

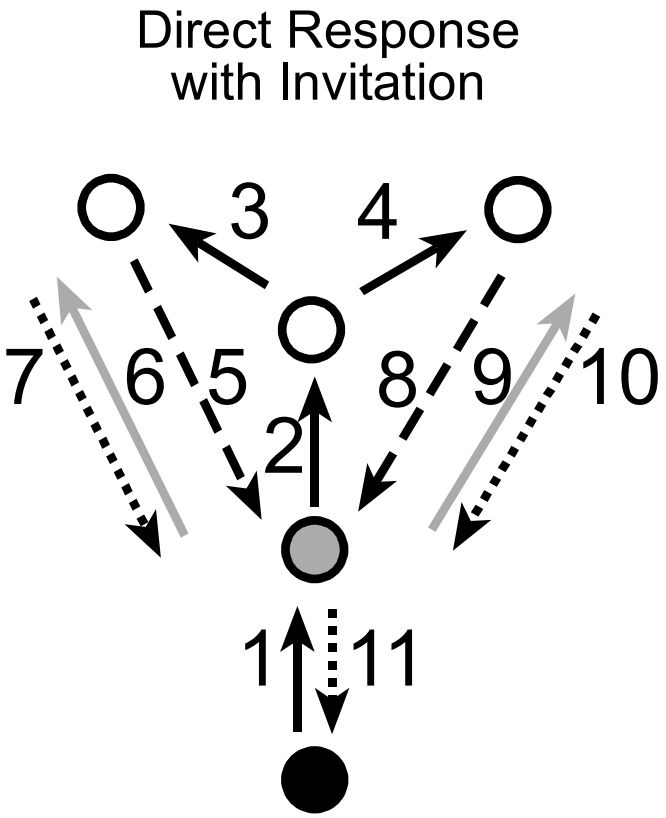
- **Familiar and scalable I/O abstractions**
  - **ala sequential, stateful, segmented file I/O**
  - ***"Open file; multiple seq. reads; close file"***
  - ***"Open transaction; multiple seq. receives; close transaction"***
  - **Synchronous and asynchronous**

CRD

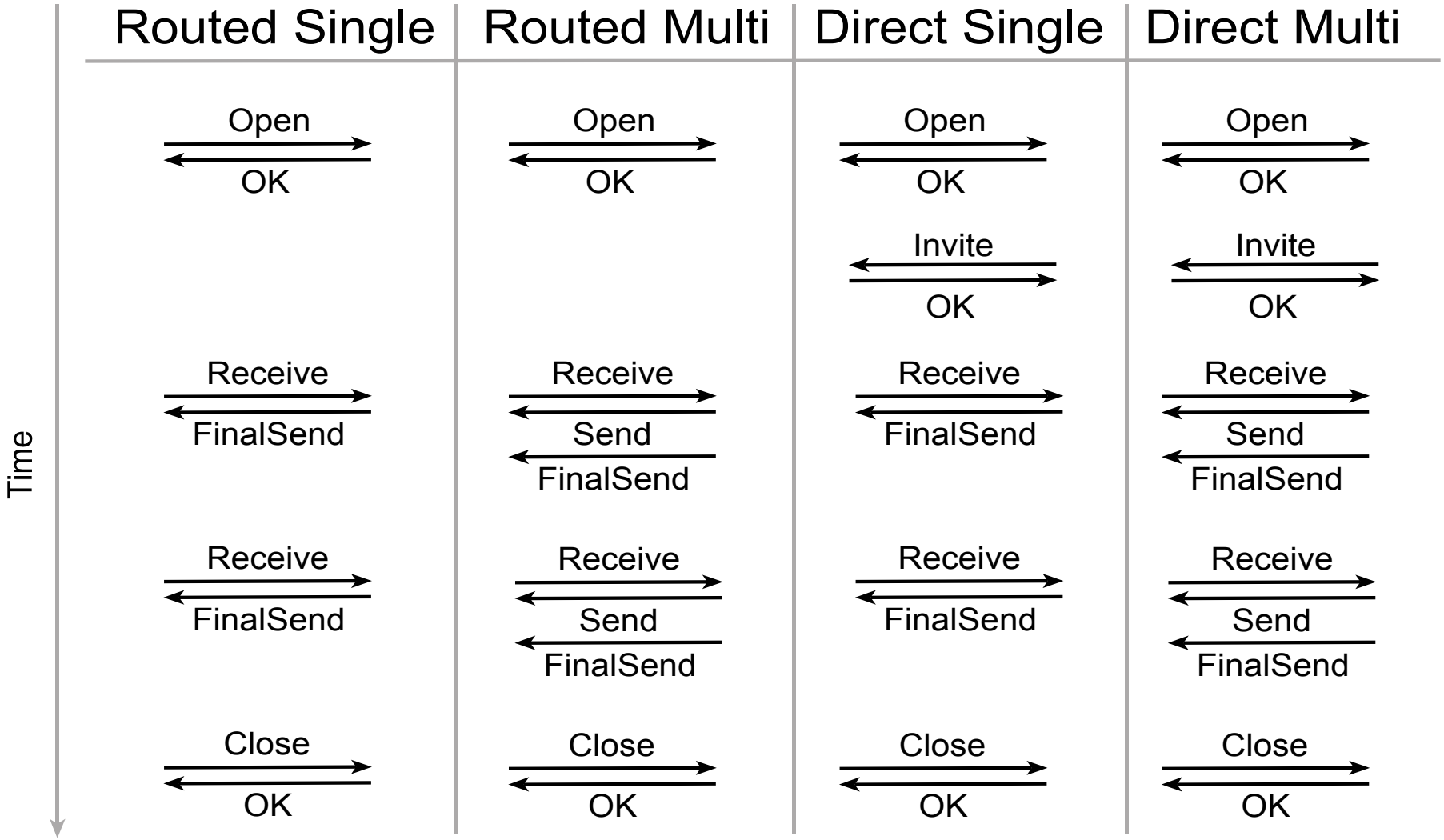
Routed vs. Direct Response Mode







CRD Message Exchange Patterns



Smart Dynamic Routing

- Allow a client to select the data and peers the query should be applied to
- Client can specify arbitrary scope and neighbor selection policies
- E.g. powerful XQuery for neighbor selection

Status

- P2PIO specification v 0.5
  http://dsd.lbl.gov/firefish/p2pio-spec/spec.pdf
- Firefish infrastructure
  - Engine implements subset of P2PIO
  - XQuery, SQL, XPath, and regular expression support for queries
  - Available soon (license pending)
    http://dsd.lbl.gov/firefish/

Outline

- DSD Research Overview (brief)
- Reliable and Secure Group Communication
- Scalable and Secure P2P Information Sharing
  - Security challenges
  - Scalability challenges
- Grid and P2P Convergence?

Grid and P2P Convergence?

- Grid or P2P?
  - Large scale
  - Heterogeneity
  - Lack of central control
  - Unreliable components
  - Frequent dynamic change

Grid and P2P Convergence?

- Grid but not P2P?
  - Security model
    - Accountability
    - Policy and trust
  - Infrastructure
    - Reusable components
    - Standards

Security Challenges

- Group security - SGL
- Authorization in P2P – scishare, Akenti
- Message-based (end-to-end) security - CryptoGrid
- Flexible security – PCCE

Need Reusable Components and Open Standards in P2P

- Resource Discovery
- Group Management
- …

- GGF Peer-to-Peer Area
- IRTF Peer-to-Peer Research Group
- ?